

1/5/2 (Item 2 from file: 351)  
DIALOG(R) File 351:Derwent WPI  
(c) 2006 The Thomson Corporation. All rts. reserv.

0013104875 - Drawing available  
WPI ACC NO: 2003-186319/200319  
XRPX Acc No: N2003-146837

**Virtual address data transfer system for multi-processor computer network requests pre-read of transmitting device translation lookaside buffer**

Patent Assignee: NEC CORP (NIDE)  
Inventor: SHINOMIYA K; SHINOMIYA K N C; SHINOMIYA N C

**Patent Family** (8 patents, 32 countries)

Patent Number	Kind	Date	Application Number	Kind	Date	Update
EP 1283473	A2	20030212	EP 200217644	A	20020806	200319 B
JP 2003050743	A	20030221	JP 2001239590	A	20010807	200323 E
US 20030033431	A1	20030213	US 2002210052	A	20020802	200325 E
JP 3594082	B2	20041124	JP 2001239590	A	20010807	200477 E
US 6928529	B2	20050809	US 2002210052	A	20020802	200552 E
EP 1283473	B1	20060503	EP 200217644	A	20020806	200635 E
DE 60211076	E	20060608	DE 60211076	A	20020806	200638 E
			EP 200217644	A	20020806	
DE 60211076	T2	20060921	DE 60211076	A	20020806	200662 E
			EP 200217644	A	20020806	

Priority Applications (no., kind, date): JP 2001239590 A 20010807; EP 200217644 A 20020806

**Patent Details**

Number	Kind	Lan	Pg	Dwg	Filing Notes
EP 1283473	A2	EN	28	13	
Regional Designated States, Original: AL AT BE BG CH CY CZ DE DK EE ES FI					
FR GB GR IE IT LI LT LU LV MC MK NL PT RO SE SI SK TR					
JP 2003050743	A	JA	20		
JP 3594082	B2	JA	24		Previously issued patent JP 2003050743

EP 1283473	B1	EN			
Regional Designated States, Original: DE FR GB					
DE 60211076	E	DE			Application EP 200217644
					Based on OPI patent EP 1283473
DE 60211076	T2	DE			Application EP 200217644
					Based on OPI patent EP 1283473

**Alerting Abstract EP A2**

NOVELTY - A request to pre-read the translation lookaside buffer (TLB) of computer 1 (700a) transmitting device, which holds recent address translation data, is transmitted by computer 1 in advance before transfer data is received by computer 2 (700b), causing virtual-address translation data to be pre-registered in the receiving TLB of computer 2

USE - For data transfer between virtual addresses within multi-processor memory-sharing computer systems.

ADVANTAGE - The apparent address translation time is shortened to make the data transfer between virtual addresses equivalent to that between real addresses, and the capacity of a data standby buffer within the destination computer can be reduced. The possibility of performance decline due to TLB misses and the switched-network bandwidth are both reduced.

DESCRIPTION OF DRAWINGS - 700a Computer 1  
700b Computer 2

**Title Terms/Index Terms/Additional Words:** VIRTUAL; ADDRESS; DATA; TRANSFER; SYSTEM; MULTI; PROCESSOR; COMPUTER; NETWORK; REQUEST; PRE; READ; TRANSMIT ; DEVICE; TRANSLATION; BUFFER

**Class Codes**

International Classification (Main): G06F-012/10  
(Additional/Secondary): G06F-012/08, G06F-015/16, G06F-015/177  
International Classification (+ Attributes)  
IPC + Level Value Position Status Version  
G06F-0012/10 A I F B 20060101  
G06F-0012/10 A I F 20060101

G06F-0012/10 A I R 20060101  
G06F-0012/10 C I R 20060101  
US Classification, Issued: 709245000, 711203000, 711154000, 711200000,  
711206000, 709245000

File Segment: EPI;

DWPI Class: T01

Manual Codes (EPI/S-X): T01-F05E; T01-H03A; T01-N01D

(19)日本国特許庁 (J P)

(12) 公 開 特 許 公 報 (A)

(11)特許出願公開番号

特開2003-50743

(P2003-50743A)

(43)公開日 平成15年2月21日(2003.2.21)

(51)Int.Cl.<sup>7</sup>

識別記号

F I

テームコード\*(参考)

G 0 6 F 12/10

5 0 1

G 0 6 F 12/10

5 0 1 Z

5 B 0 0 5

5 0 9

5 0 9 Z

5 B 0 4 5

5 5 1

5 5 1 Z

5 5 7

5 5 7

15/16

6 4 5

15/16

6 4 5

審査請求 有

請求項の数 7 O L (全 20 頁)

(21)出願番号 特願2001-239590(P2001-239590)

(22)出願日 平成13年8月7日(2001.8.7)

(71)出願人 000004237

日本電気株式会社

東京都港区芝五丁目7番1号

(72)発明者 四宮 潔

東京都港区芝五丁目7番1号 日本電気株式会社内

(74)代理人 100088328

弁理士 金田 暢之 (外2名)

Fターム(参考) 5B005 JJ11 KK13 KK22 LL11 MM21

MM51 NN22 NN71 RR01

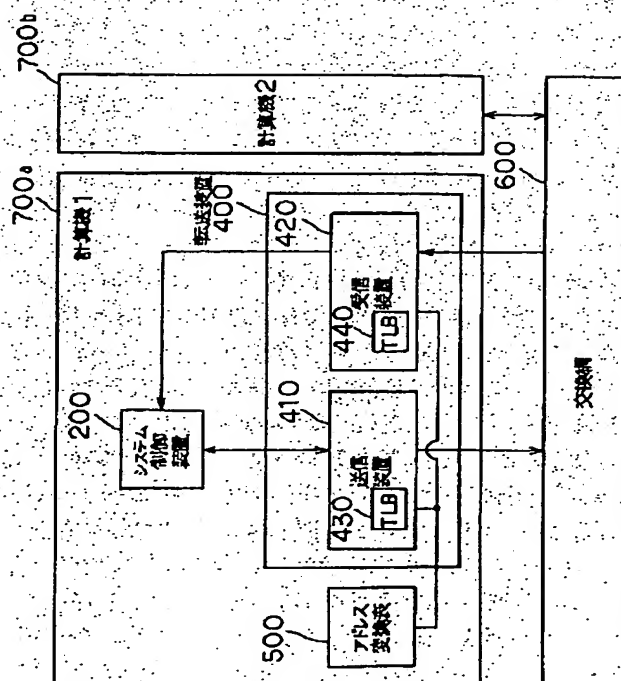
5B045 DD07 DD11

(54)【発明の名称】 仮想アドレス間データ転送方式

(57)【要約】

【課題】 アドレス変換を高速に行うとともに、データ転送開始時及びページ境界越え時にTLBミスが発生する可能性を低減する。

【解決手段】 計算機1から計算機2へのデータ転送を計算機1が命令する際、計算機1の送信装置410は、最初のデータ或いは転送先仮想アドレスがページ境界越えの最初のデータの読み出しと平行して、計算機2に転送先仮想アドレスを含むTLB先読みを指示し、計算機2のアドレス変換表500からTLB440に仮想アドレス変換情報を事前登録させる。計算機1から計算機2へのデータ転送を計算機2が命令する際、計算機1の受信装置420は、最初のデータ或いは転送先仮想アドレスがページ境界越えの最初のデータの読み出しと平行して、計算機2に転送先仮想アドレスを含むTLB先読みを指示し、計算機2のアドレス変換表500からTLB440に仮想アドレス変換情報を事前登録させる。



## 【特許請求の範囲】

【請求項 1】 中央処理装置と、データを記憶する主記憶装置と、仮想アドレスを実アドレスに変換するための変換情報を保持するアドレス変換表と、前記アドレス変換表の変換情報の一部を保持する T L B が組み込まれた前記主記憶装置のデータを読み出して複数の計算機を接続する交換網に送信する送信装置と、前記アドレス変換表の変換情報の一部を保持する T L B が組み込まれた前記交換網からのデータを受信して前記主記憶装置に書き込む受信装置とが実装された複数の計算機を有し、前記複数の計算機間で、データ転送元及び転送先のアドレスを前記仮想アドレスで指定して前記主記憶装置のデータを前記交換網を介して互いに転送するとともに、前記計算機内の前記送信装置及び前記受信装置の内部で、データ転送元及び転送先の仮想アドレスに対応する変換情報を前記アドレス変換表から前記 T L B に登録し、該 T L B に登録した変換情報を用いて前記仮想アドレスから前記実アドレスへのアドレス変換を行う仮想アドレス間データ転送方式において、

前記送信装置は、当該送信装置が実装された計算機内の主記憶装置のデータを他の計算機内の受信装置に送信する場合、当該送信装置が実装された計算機内の主記憶装置からデータを読み出すことと平行して、前記他の計算機内の受信装置に対して、データ転送先の仮想アドレスに対応する変換情報を前記他の計算機内の T L B に事前登録することを指示することを特徴とする仮想アドレス間データ転送方式。

【請求項 2】 前記受信装置は、当該受信装置が実装された計算機内の主記憶装置のデータを要求元の計算機の主記憶装置に転送するリモートリード命令を該要求元の計算機内の送信装置から指示された場合、当該受信装置が実装された計算機内の主記憶装置からのデータ読み出しと平行して、前記要求元の計算機内の受信装置に対して、データ転送先の仮想アドレスに対応する変換情報を前記他の計算機内の T L B に事前登録することを指示することを特徴とする請求項 1 に記載の仮想アドレス間データ転送方式。

【請求項 3】 前記送信装置及び前記受信装置は、前記主記憶装置における予め決められた記憶空間の大きさであるページ毎に前記アドレス変換を行うことを特徴とする請求項 2 に記載の仮想アドレス間データ転送方式。

【請求項 4】 前記送信装置は、当該送信装置が実装された計算機内の主記憶装置のデータを他の計算機内の受信装置に送信する場合、データ送信開始時に前記他の計算機内の受信装置に対して前記 T L B への事前登録を指示し、更に、前記他の計算機内の前記主記憶装置の仮想アドレスがページ境界越えを起こした時にも、該他の計算機内の受信装置に対して前記 T L B への事前登録を指示することを特徴とする請求項 3 に記載の仮想アドレス間データ転送方式。

【請求項 5】 前記送信装置は、他の計算機の主記憶装置内のデータを前記送信装置が実装された計算機内の主記憶装置を転送先として転送するために、他の計算機に転送要求を送信する際、転送要求元である前記転送先の仮想アドレスがページ境界越えを起こした場合に、他の計算機の受信装置に対してアドレス変換情報の事前登録を指示するフラグをパケットに付加した転送要求を送信し、前記他の計算機の受信装置は、前記先読みを指示するフラグがパケットに付加されている場合にのみ、前記転送要求を送信した計算機の受信装置に対して前記 T L B への事前登録を指示するパケットを送信することを特徴とする請求項 3 または請求項 4 に記載の仮想アドレス間データ転送方式。

【請求項 6】 前記送信装置は、前記 T L B への事前登録を指示する際に、データ転送先の仮想アドレスを含む T L B 先読みパケットを送信することを特徴とする請求項 1 乃至 5 のいずれか 1 項に記載の仮想アドレス間データ転送方式。

【請求項 7】 前記受信装置は、前記 T L B への事前登録を指示する際に、データ転送先の仮想アドレスを含む T L B 先読みパケットを送信することを特徴とする請求項 2 乃至 6 のいずれか 1 項に記載の仮想アドレス間データ転送方式。

## 【発明の詳細な説明】

## 【0001】

【発明の属する技術分野】本発明は、データ転送元及び転送先のアドレスを仮想アドレスで指定して互いにデータ転送を行う複数の計算機を有する仮想アドレス間データ転送方式に関し、特に、各計算機内の送信装置及び受信装置が、仮想アドレスを実アドレスに変換するための変換情報の一部を保持する T L B をそれぞれ具備し、前記 T L B に保持された変換情報を用いて仮想アドレスを実アドレスに変換する仮想アドレス間データ転送方式に関する。

## 【0002】

【従来の技術】近年、単一のプロセッサの処理能力を超える処理が計算機システムに要求され、更に、計算機システムに要求される処理能力の向上分が単一のプロセッサの性能向上分をも上回る傾向にある。

【0003】このため、計算機システムとして、複数の CPU を使用し、複数の CPU でメモリ空間を共有するメモリ共有型のマルチプロセッサ装置が、より広範囲に使用されつつある。

【0004】しかしながら、メモリ共有型のマルチプロセッサ装置においては、プロセッサに組み込まれているメモリアクセス性能の向上のためのキャッシュ間の整合性、いわゆるキャッシュ・コヒーレンスを維持しようとすると、その他の性能が低下する場合があります、特に大規模のマルチプロセッサ装置でその傾向が高い。このため、メモリ共有型のマルチプロセッサ装置においては、

メモリ空間を共有するプロセッサの台数を無制限に増加できないという第1の問題がある。

【0005】前記第1の問題を解決するために、研究用途等に用いる大規模数値演算装置として、複数の計算機を交換網で接続することにより、複数の計算機でメモリ空間を共有する必要がないマルチ計算機装置が使用されている。最近では、一般の計算機においても、キャッシュの整合性を保証しない複数の計算機を交換網で接続したマルチ計算機装置を利用する機会が増えている。

【0006】しかしながら、上記のマルチ計算機装置においては、複数の計算機のそれぞれで処理したデータを別の計算機に転送する必要があるため、データ量が増加した場合には、データ処理をするためのプロセッサにおける計算機間のデータ転送に消費する時間が増加し、マルチ計算機装置全体としての処理性能が低下するという第2の問題が発生する。

【0007】前記第2の問題を解決するために、前記マルチ計算機装置においては、各計算機内にデータ転送専用の転送装置を組み込み、前記転送装置によりプロセッサが一度に処理するデータサイズよりも大きな単位でデータ転送を行うことで、転送性能を向上させるとともにプロセッサがデータ転送に消費する時間を短縮し、マルチホスト装置全体の性能を向上させる装置が開発されている。

【0008】一方、オペレーティング・システムは、初期は小型で簡単な構成であったが、最近は大型で複雑な構成となっており、ユーザー・プロセスからオペレーティング・システム・プロセス（OSプロセス）に対して、上記のデータ転送を依頼する時のプロセス間の移行のためのコンテキストスイッチのオーバーヘッドを無視することができないという第3の問題がある。

【0009】前記第3の問題を解決するために、データ転送時の転送元／転送先アドレスをユーザ・プロセスのアドレス空間である仮想アドレスで指定し、上記の転送装置にて仮想アドレスを実アドレスに変換する方式が開発されている。

【0010】前記方式によれば、データ転送指示を行う度に、毎回オペレーティングシステムに仮想アドレスを実アドレスに変換する依頼の必要性を無くすことが可能となるため、コンテキストスイッチの頻度を下げることが可能となる。

【0011】前記転送装置は、仮想アドレスを実アドレスに変換するための変換情報を保持するアドレス変換機能を備えるが、仮想アドレス全体の変換情報は巨大であるため、転送装置内に仮想アドレス全体の変換情報を保持するアドレス変換表を実装することは困難であるという第4の問題がある。

【0012】前記第4の問題を解決するため、転送装置の外部に、仮想アドレス全体の変換情報を保持するアドレス変換表を配置し、前記転送装置内に、アドレス変換

表の中から最近使用された一部の仮想アドレスに対する変換情報のみを保持する小さなアドレス変換表、すなわちTLB（Translation look aside buffer）を組み込んだ方式が広く使用されている。

【0013】

【発明が解決しようとする課題】しかしながら、TLBを使用した方式では、TLB内に必要な仮想アドレスの変換情報が保持されていない場合、すなわちTLBミスである場合は、転送装置外部のアドレス変換表から必要な変換情報を読み出す必要があるが、前記読み出し速度が低速であるため、アドレス変換を高速に行うことができないことにより、TLBミス時の性能が低下するという第5の問題がある。

【0014】さらに、データ転送開始時や、仮想アドレスの管理単位であるページ境界越えを起こした時には、TLBミスを発生する可能性が高いため、仮想アドレス間のデータ転送機能を使用する際の性能低下が発生するという第6の問題がある。

【0015】前記第5の問題及び第6の問題を解決するため、仮想アドレスを用いてデータ転送を行う仮想アドレス間転送の性能を、物理アドレスを用いてデータ転送を行う物理アドレス間転送の性能と同等にする方式が望まれている。

【0016】本発明は、前記第5の問題及び第6の問題を解決するため、転送データがデータの受信装置に届く前にアドレス変換情報の先読み要求を事前に送信することにより、アドレス変換に要する時間をデータ送信部のデータ読み出し時間の中に隠蔽し、アドレス変換を見かけ上高速に実行し、データ転送に要する時間を短縮するとともに、データ転送開始時及びページ境界越え時にTLBミスが発生する可能性を低減させる仮想アドレス間データ転送方式を提供することを目的とする。

【0017】

【課題を解決するための手段】前記目的を達成するために本発明は、中央処理装置と、データを記憶する主記憶装置と、仮想アドレスを実アドレスに変換するための変換情報を保持するアドレス変換表と、前記アドレス変換表の変換情報の一部を保持するTLBが組み込まれた前記主記憶装置のデータを読み出して交換網に送信する送信装置と、前記アドレス変換表の変換情報の一部を保持するTLBが組み込まれた前記交換網からのデータを受信して前記主記憶装置に書き込む受信装置とが実装された複数の計算機を有し、前記複数の計算機間で、データ転送元及び転送先のアドレスを前記仮想アドレスで指定して前記主記憶装置のデータを前記交換網を介して互いに転送するとともに、前記計算機内の前記送信装置及び前記受信装置の内部で、データ転送元及び転送先の仮想アドレスに対応する変換情報を前記アドレス変換表から前記TLBに登録し、該TLBに登録した変換情報を用いて前記仮想アドレスから前記実アドレスへのアドレス

変換を行う仮想アドレス間データ転送方式において、前記送信装置は、当該送信装置が実装された計算機内の主記憶装置のデータを他の計算機内の受信装置に送信する場合、当該送信装置が実装された計算機内の主記憶装置からデータ読み出しと平行して、前記他の計算機内の受信装置に対して、データ転送先の仮想アドレスに対応する変換情報を前記他の計算機内の T L B に事前登録することを指示することを特徴とする。

【0018】また、要求元となる他の計算機内の送信装置から、自計算機内の主記憶装置のデータを前記要求元計算機内の主記憶装置に転送するリモート・リード命令を受け付けた前記受信装置は、前記自計算機内の主記憶装置からのデータ読み出しと平行して、前記要求元の計算機内の受信装置に対して、データ転送先の仮想アドレスに対応する変換情報を前記要求元の計算機内の T L B に事前登録することを指示することを特徴とする。

【0019】また、前記送信装置及び前記受信装置は、前記主記憶装置における予め決められた記憶空間の大きさであるページ毎に前記アドレス変換を行うことを特徴とする。

【0020】また、前記送信装置は、当該送信装置が実装された計算機内の主記憶装置のデータを他の計算機内の受信装置に送信する場合、データ送信開始時に前記他の計算機内の受信装置に対して前記 T L B への事前登録を指示し、更に、前記他の計算機内の前記主記憶装置の仮想アドレスがページ境界越えを起こした時にも、該他の計算機内の受信装置に対して前記 T L B への事前登録を指示することを特徴とする。

【0021】また、要求元となる他の計算機内の送信装置から、自計算機内の主記憶装置のデータを前記要求元計算機内の主記憶装置に転送するリモート・リード命令を受け付けた前記受信装置は、前記要求元の計算機の仮想アドレスがページ境界越えを起こした時に、該要求元の計算機内の受信装置に対して前記 T L B への事前登録を指示することを特徴とする。

【0022】また、前記送信装置は、前記 T L B への事前登録を指示する際に、計算機間に専用線を設けて行う方式、あるいはデータ転送先の仮想アドレスを含む T L B 先読みパケットを交換網を通して前記受信装置に送信する。

【0023】また、前記受信装置は、前記 T L B への事前登録を指示する際に、計算機間に専用線を設けて行う方式、あるいはデータ転送先の仮想アドレスを含む T L B 先読みパケットを交換網を通して前記受信装置に送信する。

【0024】(作用) 本発明においては、データ送信元の計算機内の送信装置において、データ送信先の計算機内の受信装置に対してデータを送信する場合、データ送信元の計算機内の主記憶装置からのデータ読み出しと平行して、データ送信先の計算機内の受信装置に対して、

データ転送先の仮想アドレスに対応する変換情報を T L B へ事前登録することを指示することにより、データ送信元の計算機における主記憶装置からのデータ読み出しと、データ送信先の計算機におけるアドレス変換表から T L B への変換情報の読み出しとが並行して行われることになるため、データ受信装置におけるアドレス変換のためのアドレス変換表読み出しが、データ送信装置におけるデータ読み出し時間に隠蔽されるため、見かけ上のアドレス変換時間が短縮され、データ転送にかかる総所要時間が短縮される。

【0025】

【発明の実施の形態】以下に、本発明の実施の形態について図面を参照して説明する。

【0026】(第1の実施の形態) 図1は、本発明の仮想アドレス間データ転送方式の第1の実施の形態を示す図である。

【0027】図1に示すように本実施形態においては、交換網600を介して互いにデータ交換を行う計算機700a、700b(以下、それぞれ「計算機1」「計算機2」と称する場合もある)が設けられている。

【0028】以下に、計算機700aの内部構成について説明する。なお、図1においては、計算機700bの内部構成が省略されているが、計算機700bの内部構成も計算機700aと同様である。

【0029】計算機700aは、種々の命令を処理する中央処理装置であるCPU100と、種々の命令及びデータを保持する主記憶装置300と、交換網600を介して他計算機700bとの間でデータを送受信する転送装置400と、CPU100、主記憶装置300及び転送装置400に接続され、主記憶装置300へのデータの書き込み及び読み出しを行うシステム制御装置200と、転送装置400に接続され、データ転送元及び転送先の論理アドレス(仮想アドレス)を物理アドレス(実アドレス)に変換するための変換情報として実アドレスベース511(図3、図4等参照)を保持するアドレス変換表500とから構成されている。

【0030】図2を参照して転送装置400の構成について説明する。

【0031】転送装置400は、交換網600を介して他計算機700bにデータを送信する送信装置410と、他計算機700bから交換網600を介して送信されてきたデータを受信する受信装置420とから構成されている。

【0032】送信装置410及び受信装置420には、アドレス変換表500内部のアドレス変換表エントリのうち最近使用された一部のアドレス変換表エントリをアドレス変換に使用した仮想アドレスの一部である仮想アドレスベースと共に T L B エントリとして保持するアドレス変換用バッファ T L B 430、440 がそれぞれ組み込まれている。 T L B はアドレス変換用の小容量のバ



ッファを示し、Translation Look Aside Bufferの頭文字の略号である。以降ではTLBをアドレス変換バッファを示すものとして使用する。

【0033】図3を参照して転送装置400内の送信装置410の構成について説明する。

【0034】送信装置410は、システム制御装置を経由して主記憶装置300からのデータを登録するデータバッファ411と、CPU100からの転送指示をシステム制御装置100を経由して登録する転送指示処理部412と、ユーザプロセスの仮想アドレスで指定された転送元アドレスを、主記憶装置300上のアドレスである実アドレスに変換するアドレス変換用バッファTLB430と、転送指示処理部412からの仮想アドレスとTLB内の各エントリの仮想アドレスベース431を比較し、一致した場合には実アドレスベース432から実アドレスを生成して転送指示処理部412に返却し、不一致の場合にはアドレス変換表500内の対応する実アドレスベースを読み出し、前記転送指示処理部から受け取った仮想アドレスと前記アドレス変換表500から読み出した実アドレスベース511を1組のTLBエントリとしてTLBに登録し、再度TLBを索引した後に一致した仮想アドレスベース431を持つTLBエントリの実アドレスベース432から実アドレスを生成し転送指示処理部412に返却する、TLB430内に実装されているアドレス比較部433と、転送指示処理部412から指示された実アドレスの主記憶装置300上のデータをシステム制御装置200を経由して読み出す要求を発行するデータ読出部418と、転送指示処理部412からの指示とデータバッファ411から受け取ったデータからパケットを生成し、交換網600に出力するデータ出力部417とから構成されている。

【0035】データバッファ411は、主記憶装置300から読み出されたデータをシステム制御装置200を経由して受け付け、一時的に保持し、前記読み出されたデータをデータ出力部417に転送する。

【0036】転送指示処理部412は、CPU100からの転送指示をシステム制御装置200を経由して受け付け、前記転送指示を命令800に保持する。命令800は命令コード801、転送先仮想アドレス(RV)413、転送回数414(LEN)、転送元仮想アドレス(SV)415、及び宛先計算機番号(DST)416から構成される。転送指示処理部412は命令800に従ってデータ出力部417に転送パケットの生成及び交換網600への出力を指示する。

【0037】また、転送指示処理部412は、予め決められたデータ長分のデータを転送回数414で指示された回数分転送するまでの間、1パケット分のデータを転送する度に転送先仮想アドレス413及び転送元仮想アドレス415を再計算する。なお、転送先仮想アドレス413及び転送元仮想アドレス415の再計算は、あら

かじめ定められた1パケットのデータ転送長を元に再計算され、例えば1回のデータ転送長を64バイトとすると、新仮想アドレス=(旧仮想アドレス+64)により計算する。

【0038】TLB430は、転送指示処理部412内の転送元仮想アドレス415を実アドレスに変換するために、アドレス変換表500内のアドレス変換表エントリである実アドレスベース511の一部のアドレス変換表エントリと、前記アドレス変換表エントリである実アドレスベース442を読み出す際に使用した仮想アドレスの一部である仮想アドレスベースを1組とするTLBエントリとして保持する。各TLBエントリは、仮想アドレスの一部である仮想アドレス・ベース431及び実アドレスの一部である実アドレス・ベース432として保持される。

【0039】アドレス比較部433は、TLB430の仮想アドレス・ベース431と転送元仮想アドレス415の一部である仮想アドレス・ベースとを比較し、該比較結果に基づいてTLBヒット/ミス进行判定する。転送元仮想アドレス415の一部である仮想アドレス・ベースと一致するTLBエントリが仮想アドレス・ベース431に登録されていた場合をTLBヒットとし、登録されていない場合をTLBミスとする。

【0040】データ読出部418は、アドレス比較部433にてTLBヒットと判定された場合に、TLB430の実アドレス・ベース432と転送元仮想アドレス415の一部であるオフセットとから実アドレスを生成し、主記憶装置300の前記生成された実アドレスからのデータ読み出しをシステム制御装置200に要求する。

【0041】データ出力部417は、転送指示処理部412にて受け付けられた命令800に従い、データバッファ411からデータ出力部417に転送されるデータと、転送指示処理部412内の命令コード801と、転送先アドレス413と、転送元アドレス415と、宛先計算機番号416とを用いてリモート・ライト・パケット、リモート・リード・パケット及びTLB先読みパケットを生成し、交換網600に送信する。なお、リモート・ライト・パケット、リモート・リード・パケット及びTLB先読みパケットについての詳細な説明は後述する。

【0042】図3を参照して、アドレス変換表500について説明する。

【0043】アドレス変換表500は複数のアドレス変換表エントリから構成され、アドレス変換表エントリは実アドレスベース511を保持する。複数の実アドレスベース511はそれぞれの実アドレスベース番号が付けられ、送信装置410あるいは受信装置420からの実アドレスベース読み出し要求内の仮想アドレスベースを実アドレスベース番号として実アドレスベースを読み出

し、読み出した実アドレスベースを返却される。

【0044】図4を参照して転送装置400内の受信装置420の構成について説明する。

【0045】受信装置420は、交換網600からの転送されてきたパケットのデータを受け取るデータバッファ421と、交換網600から転送されてきたパケットのデータ以外の部分を受け取り、受け取ったパケット内の命令項目の内容がリモートライト命令であった場合には、宛先仮想アドレスをTLB440を使用して実アドレスに変換し、該変換した実アドレスに対してデータバッファ421が交換網600から受け取ったデータをメモリアクセス部423を経由してシステム制御装置200に主記憶装置300に書き込むことを指示し、前記受け取ったパケット内の命令項目の内容がリモートリード命令であった場合には、メモリアクセス部423に対してシステム制御装置200を経由して主記憶装置300内のデータの読み出しを指示し、前記読み出したデータを交換網出力部425を経由して交換網600に転送することを指示し、前記受け取ったパケット内の命令項目がTLB先読みであった場合には、前記受け取ったパケット内の宛先仮想アドレスを実アドレスに変換することをTLB440に指示し、前記仮想アドレスを実アドレスに変換するためのTLBエントリを事前に登録することを指示するコマンドアドレスバッファ422と、コマンドアドレスバッファ422からの指示によりデータバッファ421からのデータをシステム制御装置200を経由して主記憶装置300にライトするか、システム制御装置200を経由して主記憶装置300内のデータを読み出し、前記読み出したデータをアドレス・コマンド・バッファ422に返却するメモリアクセス部423と、コマンド・アドレスバッファ422からの指示により交換網600にパケットを出力する交換出力部425と、コマンド・アドレスバッファ422の指示により仮想アドレスを実アドレスに変換するTLB440とから構成されている。

【0046】TLB440はコマンド・アドレス・バッファ422からの指示によりTLBエントリを検索し、ヒットであればTLB内の実アドレスベース442を用いて実アドレスを生成し、TLBミスした場合、アドレス変換表500から読み出した実アドレスベースにより、実アドレスを生成し、前記生成した実アドレスをコマンドアドレスバッファ422に返却するアドレス比較部443と、仮想アドレスベース441と実アドレスベース442を1組のTLBエントリとし、複数あるいは単数のTLBエントリを保持するTLBエントリ配列444とから構成される。

【0047】データバッファ421は、他計算機700bから交換網600を介して送信されてきたデータを受信し、一時的に保持し、前記受信データをメモリアクセス部423に転送する。

【0048】TLB440は、コマンドアドレスバッファ422に受信された転送先仮想アドレスを実アドレスに変換するために、アドレス変換表500におけるエントリの一部を保持する。各エントリは、仮想アドレスの一部である仮想アドレス・ベース441及び実アドレスの一部である実アドレス・ベース442として保持される。

【0049】アドレス比較部443は、TLB440の仮想アドレス・ベース441とコマンドアドレスバッファ422に受信された転送先仮想アドレスの一部である仮想アドレス・ベースとを比較し、その比較結果に基づいてTLBヒット／ミスを判定する。コマンドアドレスバッファ422に受信された転送先仮想アドレスの一部である仮想アドレス・ベースと一致するエントリが仮想アドレス・ベース441に存在した場合をTLBヒットとし、存在しない場合をTLBミスとする。

【0050】メモリアクセス部423は、アドレス比較部443にてTLBヒットと判定された場合に、TLB440の実アドレス・ベース442とコマンドアドレスバッファ422に受信された転送先仮想アドレスの一部であるオフセットとから実アドレスを生成し、主記憶装置300の実アドレスへのデータ書き込みをシステム制御装置200に要求する。

【0051】交換網出力部425は、コマンドアドレスバッファ422に受信された転送先仮想アドレス等を用いてリモート・ライト・パケット及びTLB先読みパケットを生成し、交換網600に送信する。なお、リモート・ライト・パケット及びTLB先読みパケットについての詳細な説明は後述する。

【0052】図5を参照して、本実施形態における仮想アドレスを実アドレスに変換するアドレス変換動作を説明する。

【0053】仮想アドレスを実アドレスに変換する場合は、送信装置410及び受信装置420にそれぞれ組み込まれているTLB430、440を用いる。

【0054】TLB430、440は、TLBエントリ配列434、444を持ち、複数あるいは単数のTLBエントリを仮想アドレス・ベース431、432、実アドレス・ベース441、442として保持している。

【0055】図5では仮想アドレスがビット31からビット13までの19ビットの仮想アドレスベースとビット12からビット0までの13ビットのオフセットで構成されているが、前記ビット数は説明のためであり、仮想アドレスベースとオフセットのビット数及び仮想アドレス全体のビット数は任意のビット数を使用することができる。

【0056】TLB索引時には、全てのTLBエントリの仮想アドレス・ベース0～Nと、実アドレスに変換しようとする変換対象の仮想アドレスの仮想アドレス・ベースとが比較され、前記変換対象の仮想アドレスの仮想



アドレス・ベースと一致するTLBエントリが仮想アドレス・ベース0～Nのいずれかに存在した場合をTLBヒットとし、存在しない場合をTLBミスとする。

【0057】TLBヒットの場合は、上記の一致したTLBエントリ内の実アドレス・ベースを変換後の実アドレスの実アドレス・ベースとし、前記変換対象の仮想アドレスの下位13ビットのオフセットを変換せずにそのまま上記の実アドレス・ベースと結合したものを変換後の実アドレスとする。

【0058】一方、TLBミスの場合は、アドレス変換表500内のアドレス変換表エントリの中から前記変換対象仮想アドレスの仮想アドレスベースの値を実アドレスベース番号とし、アドレス変換表エントリを読み出し、前記読み出したアドレス変換表エントリの実アドレスベース511をTLB430、440のTLBエントリに登録した後、再度TLB430、440を索引して前記アドレス変換を行う。前記TLBエントリ登録後のアドレス再度のアドレス変換は、エントリ登録済みであるため、必ずTLBヒットとなる。

【0059】次に、図6を参照して、アドレス変換表500におけるアドレス変換表エントリをTLB430、440のTLBエントリに登録する動作について説明する。

【0060】アドレス変換表500におけるアドレス変換表エントリ（図6では「実アドレス・ベース1～N」）を変換対象の仮想アドレスの仮想アドレス・ベースを実アドレスベース番号として索引し、前記仮想アドレス・ベースに対応する実アドレス・ベース511を読み出す。

【0061】続いて、TLB430、440のTLBエントリからLRU等の予め決められた手順でTLBエントリを選択し、前記選択されたTLBエントリに、前記読み出された実アドレス・ベースと該実アドレス・ベースの読み出しに用いた仮想アドレス・ベースとを1組のTLBエントリとして登録する。本TLBエントリ登録動作により、TLB430、440の値が更新される。

【0062】なお、登録TLBエントリの選択方法としては、ランダム、LRU (Least Recently Used)、或いはラウンドロビン等の幾つかの方法を用いることができる。

【0063】また、アドレス変換表500は、転送装置400を使用するユーザ・プロセスによる初期化時に、図示しない経路で更新される。

【0064】次に、図7を参照して、自計算機のデータを他計算機に書き込むリモート・ライト動作について説明する。

【0065】図7は計算機1のデータを計算機2に書き込む場合のリモート・ライト動作を示す。

【0066】まず、計算機1のCPU100及び計算機2のCPU100は、計算機1、計算機2内のそれぞれ

のアドレス変換表500に、データ転送を命令する各プロセスのアドレス変換に必要となるアドレス変換表500内の実アドレス・ベース511を設定する（順序1）。

【0067】次に、計算機1のCPU100は、送信装置410に対してリモート・ライト命令を、転送指示装置412内の命令800に、命令コード801、転送元仮想アドレス415、転送先仮想アドレス413、及びデータ転送長414、宛先計算機番号416として設定する（順序2）。

【0068】次に、計算機1の送信装置410は、順序2で設定された転送先仮想アドレスを仮想アドレスとして含むTLB先読みパケットを生成し、計算機2の受信装置420へ転送することにより、計算機2の受信装置420にTLB先読みを要求する（順序3）。

【0069】次に、計算機1の送信装置410は、順序2で設定された転送元仮想アドレスをTLB430を使用して転送元実アドレスに変換する（順序4）また、順序4においては、計算機2の受信装置420は、順序3で計算機1の送信装置410から転送されてきたTLB先読みパケットを受信し、前記TLB先読みパケット内の仮想アドレスでTLB440を索引して仮想アドレスを実アドレスに変換する。

【0070】順序4におけるTLB索引がTLBミスである場合、計算機1の送信装置410はアドレス変換表500の実アドレス・ベース511の中から転送元仮想アドレスに対応する実アドレス・ベースを読み出し、該読み出した実アドレス・ベースと前記転送元仮想アドレスの仮想アドレス・ベースとを1組のTLBエントリとしてTLB430に登録し、再度TLB430を索引してアドレス変換を行う（順序5）。TLBエントリ登録後のTLB再索引は、必要な仮想アドレス・ベースを登録した直後に行われるため、TLBヒットとなり、アドレス変換が完了する。

【0071】また、順序5においては、計算機2の受信装置420は、順序4におけるTLB索引がTLBミスである場合、アドレス変換表500の実アドレス・ベース511の中から仮想アドレスに対応する実アドレス・ベースを読み出し、読み出した実アドレス・ベースと仮想アドレスの仮想アドレス・ベースとを1組のTLBエントリとしてTLB440に登録し、再度TLB440を索引してアドレス変換を行う。前記TLB再索引は、必要な仮想アドレス・ベースを登録した直後に行われるため、TLBヒットとなり、アドレス変換が完了する。

【0072】次に、計算機1の送信装置410は、システム制御装置200を経由して主記憶装置300の転送元実アドレスから書き込みデータを読み出す（順序6）。

【0073】次に、計算機1の送信装置410は、順序2で設定された転送先仮想アドレスと、順序6で読み出

した書き込みデータとを含むリモート・ライト・パケットを生成し、計算機2の受信装置420へ転送する(順序7)。

【0074】次に、計算機2の受信装置420は、順序7で計算機1の送信装置410から転送されたリモート・ライト・パケットを受信し、該リモート・ライト・パケット内の転送先仮想アドレスでTLB440を索引し、転送先仮想アドレスを転送先実アドレスに変換する(順序8)。

【0075】順序8におけるTLB索引がTLBミスである場合、計算機2の受信装置420はアドレス変換表500の実アドレス・ベース511の中から転送先仮想アドレスに対応する実アドレス・ベースを読み出し、読み出した実アドレス・ベースと転送元仮想アドレスの仮想アドレス・ベースとを1組のTLBエントリとしてTLB440に登録し、再度TLB440を索引してアドレス変換を行う。前記TLB再索引は、必要な仮想アドレス・ベースを登録した直後に行われるため、TLBヒットとなり、アドレス変換が完了する(順序9)。

【0076】順序7で受信されたリモート・ライト・パケット内の書き込みデータを計算機2の受信装置420がシステム制御装置200を経由して主記憶装置300の転送先実アドレスに書き込む(順序10)。

【0077】なお、計算機1の送信装置410と計算機2の受信装置420との間では、コマンド及びデータがパケット単位で転送される。

【0078】図8を参照して、計算機1、2間で転送されるパケットの形式を説明する。

【0079】本実施形態においては、リモート・リード・パケットと、リモート・ライト・パケットと、TLB先読みパケットとの3つのパケット形式を定義する。

【0080】図8(a)は、リモート・ライト・パケットのパケット形式を示す。

【0081】リモート・ライト・パケットは、命令項目の命令がWrite、命令項目のデータ長がデータ項目の個数である4、計算機項目が本パケットの送信(発信)計算機番号及び受信(宛先)計算機番号、アドレス項目が転送先仮想アドレス、データ項目が4つの転送データである。データ項目の数は転送するデータ項目の個数によって代わり、本実施形態におけるデータ項目の個数4は一例である。

【0082】図8(b)は、リモート・リード・パケットのパケット形式を示す。

【0083】リモート・リード・パケットは、命令項目の命令がRead、命令項目のデータ長が1、計算機項目が本パケットの送信(発信)計算機番号及び受信(宛先)計算機番号、アドレス項目が転送元仮想アドレス、データ項目が転送先仮想アドレスである。

【0084】図8(c)は、TLB先読みパケットのパケット形式を示す。

【0085】TLB先読みパケットは、命令項目の命令がTLB、データ長が0、計算機項目が本パケットの送信(発信)計算機番号及び受信(宛先)計算機番号、アドレス項目が仮想アドレスであり、データ項目は存在しない。

【0086】図8(d)は、図8(a)～(c)に示した各パケットによる命令の動作を示す。

【0087】リモート・ライト・パケットは、命令発行元の計算機のデータを宛先計算機へ転送する。

【0088】また、リモート・リード・パケットは、宛先計算機のデータを命令発行元の計算機内に転送する。

【0089】また、TLB先読みパケットは、宛先計算機の受信装置に対し、仮想アドレスでTLBを索引し、TLBミスの場合はアドレス変換表500から必要な変換情報を読み出し、TLB440に登録するよう命令する。

【0090】次に、図9を参照して、他計算機からデータを読み出し、前記データを自計算機に書き込むリモート・リード動作について説明する。

【0091】図9は、計算機2からデータを読み出し、前記データを計算機1に書き込む場合のリモート・リード動作を説明する。

【0092】まず、計算機1のCPU100及び計算機2のCPU100は、計算機1、計算機2内のそれぞれの実アドレス変換表500に、各プロセスの実アドレス変換に必要な実アドレス・ベース511を設定する(順序1)。

【0093】次に、計算機1のCPU100は、送信装置410に対してリモート・リード命令を、転送指示装置412内の命令800に、命令コード801、転送元仮想アドレス415、転送先仮想アドレス413、及びデータ転送長414、宛先計算機番号416として設定する(順序2)。

【0094】次に、計算機1の送信装置410は、順序2で設定された命令800内の命令コード801、宛先計算機番号416、転送元仮想アドレス415及び転送先仮想アドレス413を含むリモート・リード・パケットを生成し、計算機2の受信装置420へ交換網600を経由して転送する(順序3)。

【0095】次に、計算機2の受信装置420は、順序3で計算機1の送信装置410から転送されてきたリモート・リード・パケットを受信し、前記リモート・リード・パケット内の転送先仮想アドレスを宛先仮想アドレスとして含むTLB先読みパケットを生成し、前記リモート・リード・パケットの送信元である計算機1へ転送する(順序4)。

【0096】次に、計算機2の受信装置420は、順序4で受信したリモート・リード・パケット内の転送元仮想アドレスでTLB440を索引して転送元仮想アドレスを転送元実アドレスに変換する(順序5)。

【0097】また、順序5においては、計算機1の受信装置420は、順序4で計算機2の受信装置420から転送されてきたTLB先読みパケットを受信し、前記TLB先読みパケット内の宛先仮想アドレスでTLB430を索引して仮想アドレスを実アドレスに変換する。

【0098】順序5におけるTLB索引がTLBミスであった場合、計算機1の受信装置420はアドレス変換表500の実アドレス・ベース511の中から転送元仮想アドレスに対応する実アドレス・ベースを読み出し、前記読み出した実アドレス・ベースと仮想アドレスの仮想アドレス・ベースとを1組のTLBエントリとしてTLB430に登録し、再度TLB430を索引してアドレス変換を行う(順序6)。前記TLB再索引は、必要な仮想アドレス・ベースを登録した直後に行われるため、TLBヒットとなり、アドレス変換が完了する。

【0099】また、順序6においては、計算機2の受信装置420は、順序5におけるTLB索引がTLBミスであった場合、アドレス変換表500の実アドレス・ベース511の中から仮想アドレスに対応する実アドレス・ベースを読み出し、読み出した実アドレス・ベースと転送元仮想アドレスの仮想アドレス・ベースとを1組のTLBエントリとしてTLB440に登録し、再度TLB440を索引してアドレス変換を行う。前記TLB再索引は、必要な仮想アドレス・ベースを登録した直後に行われるため、TLBヒットとなり、アドレス変換が完了する。

【0100】次に、計算機2の受信装置420は、システム制御装置200を経由して主記憶装置300の前記アドレス変換後の転送元実アドレスから転送データを読み出す(順序7)。

【0101】次に、計算機2の受信装置420は、順序4で受信したリモート・リード・パケット内の転送先仮想アドレスと、順序7で主記憶装置300から読み出した転送データとを含むリモート・ライト・パケットを生成し、リモート・リード・パケットの送信元である計算機1へ転送する(順序8)。

【0102】次に、計算機1の受信装置420は、順序8で計算機2の受信装置420から転送されてきたリモート・ライト・パケットを受信し、前記リモート・ライト・パケット内の転送先仮想アドレスでTLB430を索引し、転送先仮想アドレスを転送先実アドレスに変換する(順序9)。

【0103】順序9におけるTLB索引がTLBミスであった場合、計算機1の受信装置420はアドレス変換表500の実アドレス・ベース511の中から転送先仮想アドレスに対応する実アドレス・ベースを読み出し、読み出した実アドレス・ベースと転送先仮想アドレスの仮想アドレス・ベースとを1組のTLBエントリとしてTLB430に登録し、再度TLB430を索引してアドレス変換を行う(順序10)。前記TLB再索引は、

必要な仮想アドレス・ベースを登録した直後に行われるため、TLBヒットとなり、アドレス変換が完了する。

【0104】その後、計算機1の受信装置420は、順序8で受信したリモート・ライト・パケット内の読み出しデータをシステム制御装置200を経由して主記憶装置300の前記アドレス変換後の転送先実アドレスに書き込む(順序11)。

【0105】図10を参照して、送信装置410の動作について詳細に説明する。

【0106】まず、転送指示処理部412がCPU100からの転送命令800を受け付け(ステップS110)、前記転送命令800を、命令コード801、転送先仮想アドレス(RV)413、転送回数414(LN)、転送元仮想アドレス(SV)415、及び宛先計算機番号(DST)416として保持する。

【0107】次に、転送指示処理部412に受け付けられた転送指示がリモート・ライト命令であるか否かが判断され(ステップS120)、リモート・ライト命令である場合、データ出力部417は、転送指示処理部412内の転送先仮想アドレス413及び宛先計算機番号416と、自計算機の送信計算機番号とを含むTLB先読みパケットを生成し、交換網600に送信する(ステップS130)。前記TLB先読みパケットは、交換網600により宛先計算機番号416の他計算機内の受信装置420に転送される。

【0108】次に、アドレス比較部433は、転送指示処理部412内の転送元仮想アドレス415でTLB430を索引し(ステップS140)、前記索引がTLBミスであるか否かを判定する(ステップS150)。TLB索引時、アドレス比較部433は、転送元仮想アドレス415の一部である仮想アドレス・ベースと、TLB430の仮想アドレス・ベース431とを比較し、前記比較結果が全てのTLBエントリにおいて不一致であった場合はTLBミスと判定し、一致したTLBエントリがある場合はTLBヒットと判定する。

【0109】ステップS150にてTLBミスと判定されると、図6に示したように、アドレス変換表500の実アドレス・ベース511の中から転送元仮想アドレス415に対応する実アドレス・ベースが読み出され、読み出された実アドレス・ベースと転送元仮想アドレス415の一部である仮想アドレス・ベースとが、それぞれ実アドレス・ベース432及び仮想アドレス・ベース431に1組のTLBエントリとして登録される(ステップS160)。TLBエントリ登録後に再度、TLB430を索引すると、必要な仮想アドレス・ベースを登録した直後であるため、TLBヒットとなる。

【0110】ステップS150或いはステップS160にてTLBヒットと判定されると、データ読出部418は、図5に示したように、TLB430の実アドレス・ベース432のうち転送元仮想アドレス415の仮想ア

ドレス・ベースに一致した実アドレス・ベースと、転送元仮想アドレス415の一部であるオフセットとから、転送元実アドレスを生成し、システム制御装置200を経由して主記憶装置300の前記転送元実アドレスから転送データを読み出し、データバッファ411に登録する(ステップS170)。

【0111】次に、データ出力部417は、データバッファ411内の転送データと、転送指示処理部412内の転送先仮想アドレス413及び宛先計算機番号416と、自計算機の送信計算機番号とを含むリモート・ライト・パケットを生成し、交換網600に送信する(ステップS180)。前記リモート・ライト・パケットは、交換網600を経由して宛先計算機番号416の他計算機内の受信装置420に転送される。

【0112】ステップS180における転送が終了すると、転送指示処理部412は、転送回数414を1減算して転送回数414を更新し(ステップS190)、更新した転送回数414が0よりも大きいかなかを判断する(ステップS200)。

【0113】ステップS200にて転送回数414が0である場合は転送命令を終了し、また、転送回数414が0よりも大きい場合、すなわち転送命令を継続する場合は、転送先仮想アドレス413及び転送元仮想アドレス415を1パケットのデータ長を加算することで更新した後(ステップS210)、ステップS140における処理に戻る。例えば、1パケットのデータ長が16バイトの場合、新仮想アドレス=(旧仮想アドレス+16)により更新する。

【0114】一方、ステップS120にて転送指示処理部412に受け付けられた転送命令がリモート・ライト命令ではなく、リモート・リード命令であると判断された場合、データ出力部417は、転送指示処理部412内の転送先仮想アドレス413、転送元仮想アドレス415及び宛先計算機番号416と、自計算機の送信計算機番号とを含むリモート・リード・パケットを生成し、交換網600に送信する(ステップS300)。前記リモート・リード・パケットは、交換網600により宛先ホスト番号416の他計算機内の受信装置420に転送される。

【0115】ステップS300における転送が終了すると、転送指示処理部412は、転送回数414を1減算して転送回数414を更新し(ステップS310)、更新した転送回数414が0よりも大きいかなかを判断する(ステップS320)。

【0116】ステップS320にて転送回数414が0である場合は転送命令を終了し、また、転送回数414が0よりも大きい場合、すなわち転送命令を継続する場合は、転送先仮想アドレス413及び転送元仮想アドレス415を1パケットのデータ長を加算することで更新した後(ステップS330)、ステップS300にお

る処理に戻る。例えば、1パケットのデータ長が16バイトの場合、新仮想アドレス=(旧仮想アドレス+16)により更新する。

【0117】図11を参照して、受信装置420の動作について詳細に説明する。

【0118】まず、コマンドアドレスバッファ422が交換網600からパケットを受信する(ステップR110)。

【0119】次に、コマンドアドレスバッファ422に受信されたパケットがリモート・リード・パケットであるかなが判断され(ステップR120)、リモート・リード・パケットである場合、交換網出力部425は、リモート・リード・パケット内の発信元計算機番号及び転送先仮想アドレスをそれぞれ宛先計算機番号及び仮想アドレスとして含むTLB先読みパケットを生成し、交換網600に送信する(ステップR130)。前記TLB先読みパケットは、交換網600により宛先計算機番号の他計算機内の受信装置420に転送される。

【0120】次に、アドレス比較部443は、リモート・リード・パケット内の転送元仮想アドレスでTLB440を索引し(ステップR140)、前記索引がTLBミスであるかなかを判定する(ステップR150)。ステップR150のTLB索引時、アドレス比較部443は、リモート・リード・パケット内の転送元仮想アドレスの一部である仮想アドレス・ベースと、TLB440の仮想アドレス・ベース441とを比較し、前記比較結果が全てのTLBエントリにおいて不一致であった場合はTLBミスと判定し、一致したTLBエントリがある場合はTLBヒットと判定する。

【0121】ステップR150にてTLBミスと判定された場合は、図6に示したように、アドレス変換表500の実アドレス・ベース511の中からリモート・リード・パケット内の転送元仮想アドレスに対応する実アドレス・ベースが読み出され、読み出された実アドレス・ベースとリモート・リード・パケット内の転送元仮想アドレスに対応する仮想アドレスの一部である仮想アドレス・ベースとが、それぞれ実アドレス・ベース442及び仮想アドレス・ベース441に1組のTLBエントリとして登録される(ステップR160)。前記TLBエントリ登録後、再度、TLB440を索引すると、必要な仮想アドレス・ベースを登録した直後であるため、TLBヒットとなる。

【0122】ステップR150或いはステップR160にてTLBヒットと判定されると、メモリアクセス部423は、図5に示したように、TLB440の実アドレス・ベース442のうちリモート・リード・パケット内の転送元仮想アドレスの仮想アドレス・ベースに一致した実アドレス・ベースと、リモート・リード・パケット内の転送元仮想アドレスの一部であるオフセットとから、転送元実アドレスを生成し、システム制御装置20

0を経由して主記憶装置300の前記転送元実アドレスから転送データを読み出す(ステップR170)。

【0123】交換網出力部425は、主記憶装置300から読み出した転送データと、宛先計算機番号(リモート・リード・パケット内の発信計算機番号)と、リモート・リード・パケット内の転送先仮想アドレスとを含むリモート・ライト・パケットを生成し、交換網600に送信する(ステップR180)。前記リモート・ライト・パケットは、交換網600により宛先計算機番号の計算機内の受信装置420に転送される。

【0124】以上で、リモート・リード・パケットの処理が完了する。

【0125】一方、ステップR120にてコマンドアドレスバッファ422に受信されたパケットがリモート・リード・パケット以外であると判断された場合、そのパケットがリモート・ライト・パケットであるか否かが判断される(ステップR300)。

【0126】ステップR300にてコマンドアドレスバッファ422に受信されたパケットがリモート・ライト・パケットであると判断された場合、アドレス比較部443は、リモート・ライト・パケット内の転送先仮想アドレスでTLB440を索引し(ステップR310)、前記索引がTLBミスであるか否かを判定する(ステップR320)。前記TLB索引時、アドレス比較部443は、リモート・ライト・パケット内の転送先仮想アドレスの一部である仮想アドレス・ベースと、TLB440の仮想アドレス・ベース441とを比較し、前記比較結果が全てのTLBエントリにおいて不一致であった場合はTLBミスと判定し、一致したエントリがある場合はTLBヒットと判定する。

【0127】ステップR320にてTLBミスと判定された場合は、図6に示したように、アドレス変換表500の実アドレス・ベース511の中からリモート・ライト・パケット内の転送先仮想アドレスに対応する実アドレス・ベースが読み出され、読み出された実アドレス・ベースとリモート・ライト・パケット内の転送先仮想アドレスの一部である仮想アドレス・ベースとがそれぞれ実アドレス・ベース442及び仮想アドレス・ベース441に1組のTLBエントリとして登録される(ステップR330)。TLB440の再索引では、必要な仮想アドレス・ベースを登録した直後であるため、TLBヒットとなる。

【0128】ステップR320或いはステップR330にてTLBヒットと判定されると、メモリアクセス部423は、TLB440の実アドレス・ベース442のうちリモート・ライト・パケット内の転送先仮想アドレスの仮想アドレス・ベースに一致した実アドレス・ベースと、リモート・ライト・パケット内の転送先仮想アドレスの一部であるオフセットとから、図5に示したように転送先実アドレスを生成し、システム制御装置200を

経由して主記憶装置300の前記転送先実アドレスにデータバッファ421内の転送データをメモリアクセス部423に転送し、メモリアクセス部423はシステム制御装置200を経由して主記憶装置300に書き込む(ステップR340)。

【0129】以上で、リモート・ライト・パケットの処理が完了する。

【0130】一方、ステップR300にてコマンドアドレスバッファ422に受信されたパケットがリモート・ライト・パケットでなく、TLB先読みパケットであると判断された場合、アドレス比較部443は、TLB先読みパケット内の仮想アドレスでTLB440を索引し(ステップR400)、前記索引がTLBミスであるか否かを判定する(ステップR410)。前記TLB索引時、アドレス比較部443は、TLB先読みパケット内の仮想アドレスの仮想アドレス・ベースと、TLB440の仮想アドレス・ベース441とを比較し、前記比較結果が全てのTLBエントリにおいて不一致であった場合はTLBミスと判定し、一致したエントリがある場合はTLBヒットと判定する。

【0131】ステップR410にてTLBミスと判定された場合は、図6に示したように、アドレス変換表500の実アドレス・ベース511の中からTLB先読みパケット内の仮想アドレスに対応する実アドレス・ベースが読み出され、読み出された実アドレス・ベースとTLB先読みパケット内の仮想アドレスの一部である仮想アドレス・ベースとが、それぞれ実アドレス・ベース442及び仮想アドレス・ベース441に1組のTLBエントリとして登録される(ステップR420)。

【0132】以上で、TLB先読みパケットの処理が完了する。

【0133】(第2の実施の形態) 上述した第1の実施形態においては、送信装置410は、CPU100からのリモート・ライト指示を受け付けた後、データ送信開始時にのみTLB先読みパケットを送信し、また、受信装置420は、全てのリモート・リード・パケットに対してTLB先読みパケットを送信している。

【0134】しかしながら、データ送信開始時にのみTLB先読みパケットを送信する構成では、アドレス変換をページと呼ばれる予め決められた主記憶空間の大きさ毎に行う場合に、データ送信中に仮想アドレス・ベースが変化する時、即ちページ境界越え時にTLBミスとなる可能性が高いため、ページ境界越え時にデータ転送速度が低下する問題がある。

【0135】また、全てのリモート・リード・パケットに対してTLB先読みパケットを送信する構成では、交換網の帯域幅の浪費、及びTLB索引頻度の無意味な増加を引き起こす問題がある。

【0136】本実施形態は、前記2つの問題を解決するために、送信装置410は、データ送信開始時、及びデ



ータ転送先の仮想アドレスがページ越えをした時に、TLB先読みパケットを送信し、また、受信装置420は、データ転送先の仮想アドレスがページ越えをした時にのみ、リモート・リード・パケットに対してTLB先読みパケットを送信する構成とすることで、ページ境界越え時の性能低下、交換網帯域幅浪費及びTLB索引頻度増の問題を解消するものである。

【0137】図12に、第1の実施形態からの相違点を明確にするために、送信装置410内の転送指示処理部412の転送先仮想アドレス計算部分のみを示す。

【0138】第1の実施形態における転送指示処理部412は、転送先仮想アドレス413と1パケットのデータ長であるパケット当たりのデータ長710とを仮想アドレス加算部711で加算し、仮想アドレス加算部711の出力を転送先仮想アドレス413としている。

【0139】これに対して、本実施形態における転送指示処理部412は、仮想アドレス加算部711の出力を転送先仮想アドレス413とするとともに、ページ境界越え検出信号712としてデータ出力部417に入力する。

【0140】すなわち、転送先仮想アドレス413がページ境界越えを起こした場合には、更新された転送先仮想アドレス413によるTLB先読み要求をページ境界越え検出信号712としてデータ出力部417に送信する。

【0141】ページ境界検出機能により、データ出力部417は、リモート・ライト命令の最初のリモート・ライト・パケット送信時、あるいは転送指示処理部412からのTLB先読み要求を受け付けた場合にも、TLB先読みパケットを交換網600を経由して宛先計算機に送信する。

【0142】また、データ出力部417は、リモート・リード・パケット送信中に、転送指示処理部412からのTLB先読み要求を受け付けた場合には、TLB先読み要求フラグをセットしたリモート・リード・パケットを送信し、また、リモート・リード・パケット送信中に、TLB先読み要求を受け付けてない場合には、TLB先読み要求フラグをセットしないリモート・リード・パケットを送信する。

【0143】TLB先読み要求フラグ付きのリモート・リード・パケットを図8(e)に示す。

【0144】図8(e)ではTLB先読み要求フラグPFが命令項目に追加されており、前記PF=1のリモート・リード・パケットを受け付けた受信装置420は、前記リモート・リード・パケット送信元にTLBの先読みパケットを送信し、前記PF=0のリモート・リード・パケットを受け付けた受信装置420は前記TLB先読みパケットを送信しない。

【0145】第2の実施形態の受信装置420の動作を第1の実施形態の相違点のみを示すフローチャートを図

13に示す。なお、図13において、「R+数字」のステップが第1の実施形態のステップを示し、「RX+数字」のステップが第2の実施形態で修正されたステップを示す。

【0146】図13を参照して、第2の実施形態における受信装置420の動作を説明する。

【0147】受信装置420がリモート・リード・パケットを受信した場合、(ステップR120)、ステップRX100において、TLB先読み要求フラグPFを確認し、前記PF=1の場合はステップR130に遷移する。ステップRX100において、前記PF=0の場合はステップR140に遷移する。

【0148】

【発明の効果】以上説明したように本発明においては、計算機間のデータを転送する場合、データ送信元の計算機内の主記憶装置からのデータ読み出しと同時に、データ送信先の計算機内の受信装置に対して、データ転送先の仮想アドレスに対応する変換情報をTLBへ事前登録することを指示する構成とし、データ送信元の計算機における主記憶装置からのデータ読み出しと、データ送信先の計算機におけるアドレス変換表からTLBへの仮想アドレス変換情報の読み出しとが並行して行うことにより、データ送信先の計算機でデータが受信されてからアドレス変換情報を読み出しTLBへの登録を行う従来方式と比較して、アドレス変換時間を見かけ上短縮する第1の効果がある。

【0149】更に、データ送信先の計算機におけるTLB読み出し時間の一部或いは全てが、データ送信元の計算機におけるデータ読み出し時間に隠蔽されることから、データ送信先の計算機内では、受信したデータを一時的に保持するためのデータバッファにてデータがアドレス変換表読み出しを待つ時間が短縮されるため、データ送信先の計算機内でのデータ待ちバッファの容量を縮小することができ、使用するハードウェアの量を削減することができる第2の効果がある。

【0150】また、データ送信元の計算機内の受信装置において、データ送信先の計算機内の送信装置からの指示に従って該データ送信先の計算機内の受信装置に対してデータを送信する場合にも、データ送信元の計算機内の主記憶装置からのデータ読み出しと同時に、データ送信先の計算機内の受信装置に対して、当該データの転送先の仮想アドレスに対応する変換情報をTLBへ事前登録することを指示する構成とすることで、上記と同様の効果が得られる第3の効果がある。

【0151】また、データ送信元の計算機内の送信装置において、データ送信先の計算機内の受信装置にデータを送信する場合、データ送信開始時にデータ送信先の計算機内の受信装置にTLBへの事前登録を指示し、更に、データ送信先の計算機内の主記憶装置の仮想アドレスがページ境界越えを起こした時にもTLBへの事前登

10

20

30

40

50



録を指示する構成とすることにより、データ送信開始時及びページ境界越え時に、TLB内に必要な仮想アドレスの変換情報が登録されていない状態、すなわちTLBミスが発生することによる性能低下の可能性を低減することができるという第4の効果がある。

【0152】また、データ送信元の計算機内の受信装置において、データ送信先の計算機内の送信装置からの指示に従って該データ送信先の計算機内の受信装置に対してデータを送信する場合、データ送信先の計算機内の主記憶装置の仮想アドレスがページ境界越えを起こした時にのみ、TLBへの事前登録を指示する構成とすることにより、TLB索引頻度の過剰な増加を抑制することができるとともに、交換網の帯域幅を低下させることができるという第5の効果がある。

#### 【図面の簡単な説明】

【図1】本発明における仮想アドレス間データ転送方式の第1の実施の形態を示す図である。

【図2】図1に示した転送装置の構成を説明するための図である。

【図3】図2に示した送信装置の構成を説明するための図である。

【図4】図2に示した受信装置の構成を説明するための図である。

【図5】図1に示した仮想アドレス間データ転送方式における、TLBを用いたアドレス変換動作を説明するための図である。

【図6】図1に示した仮想アドレス間データ転送方式における、アドレス変換表の読み出し動作を説明するための図である。

【図7】図1に示した仮想アドレス間データ転送方式におけるリモート・ライト動作を説明するための図である。

【図8】本発明の仮想アドレス間データ転送方式に用いられるパケットのフォーマット形式を説明するための図であり、(a)はリモート・ライト・パケットのフォーマット形式を示す図、(b)はリモート・リード・パケットのフォーマット形式を示す図、(c)はTLB先読みパケットのフォーマット形式を示す図、(d)は(a)～(c)に示した各パケットによる命令の動作を示す図、(e)は先読み指示フラグ付きのリモート・リードパケットを示す図である。

【図9】図1に示した仮想アドレス間データ転送方式におけるリモート・リード動作を説明するための図である。

【図10】図3に示した送信装置の動作を説明するためのフローチャートである。

【図11】図4に示した受信装置の動作を説明するためのフローチャートである。

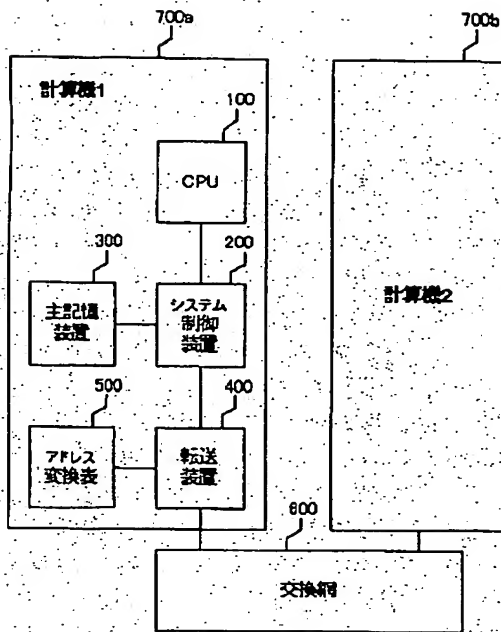
【図12】本発明の仮想アドレス間データ転送方式の第2の実施の形態に用いられる送信装置内の転送処理部を説明するための図である。

【図13】本発明の仮想アドレス間データ転送方式の第2の実施の形態に用いられる受信装置の動作を説明するためのフローチャートである。

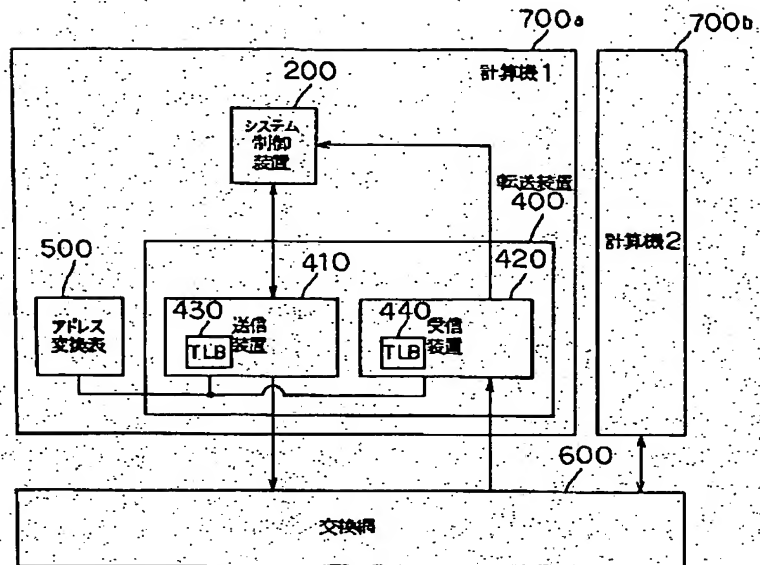
#### 【符号の説明】

100	中央処理装置
200	システム制御装置
300	主記憶装置
400	データ転送装置
410	送信装置
411	送信装置のデータバッファ
412	送信装置の転送指示処理部
413	転送先仮想アドレス
414	転送回数
415	転送元仮想アドレス
416	宛先計算機番号
417	送信装置のデータ出力部
418	送信装置のデータ読出部
433	送信装置のアドレス比較部
420	受信装置
421	受信装置のデータバッファ
422	受信装置のコマンドアドレスバッファ
423	受信装置のメモリアクセス部
443	受信装置のアドレス比較部
425	受信装置の交換網出力部
430	送信装置のTLB
431	送信装置のTLBの仮想アドレス・ベース部
432	送信装置のTLBの実アドレス・ベース部
440	受信装置のTLB
441	受信装置のTLBの仮想アドレス・ベース部
442	受信装置のTLBの実アドレス・ベース部
500	アドレス変換表
600	交換網
700a, 700b	計算機
710	データ長
711	仮想アドレス加算部
712	ページ境界越え検出信号

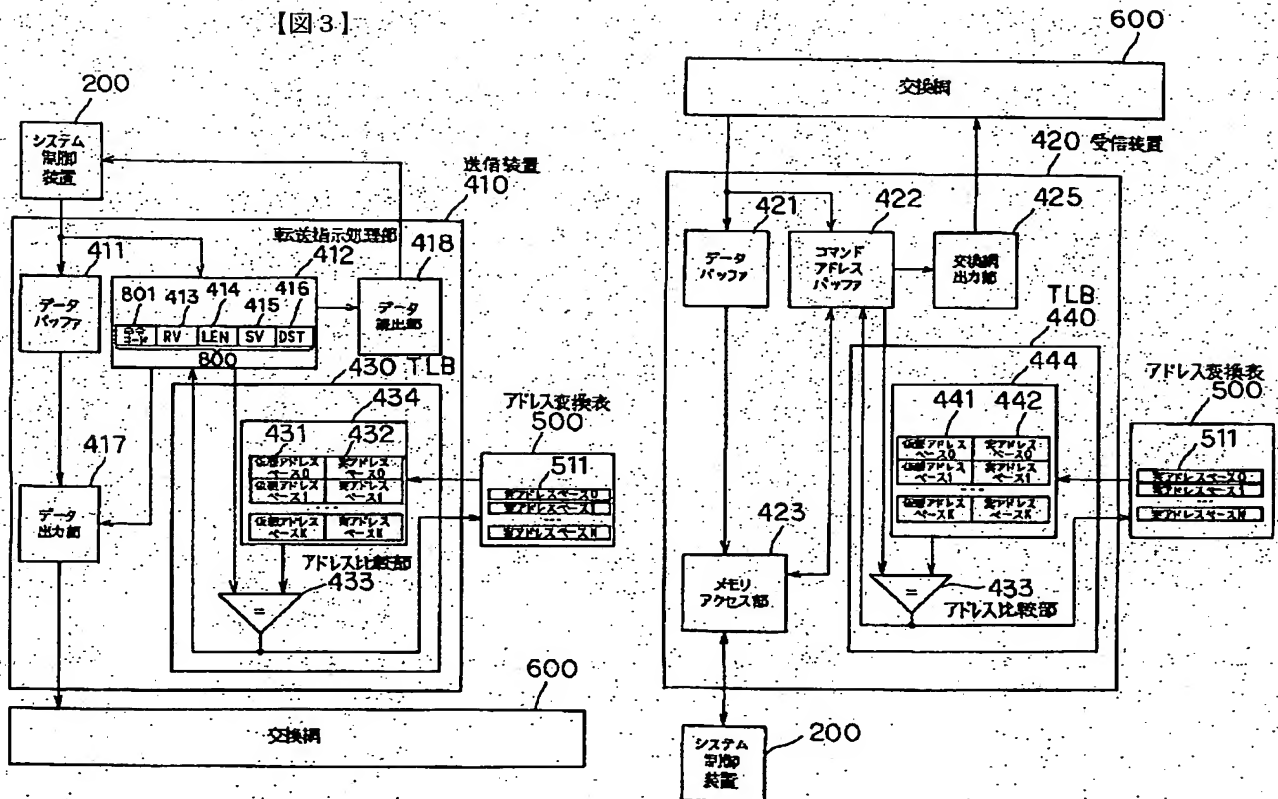
【図 1】



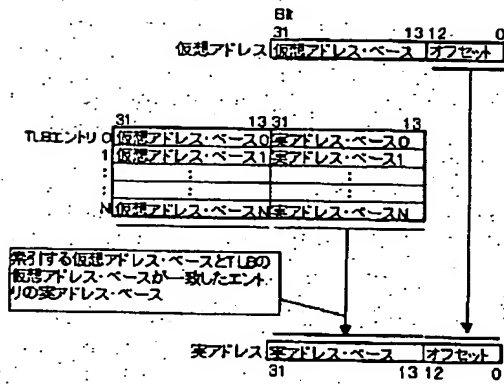
【図 2】



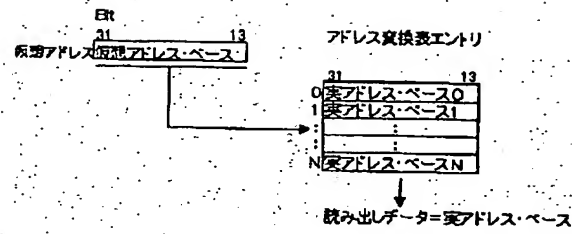
【図 4】



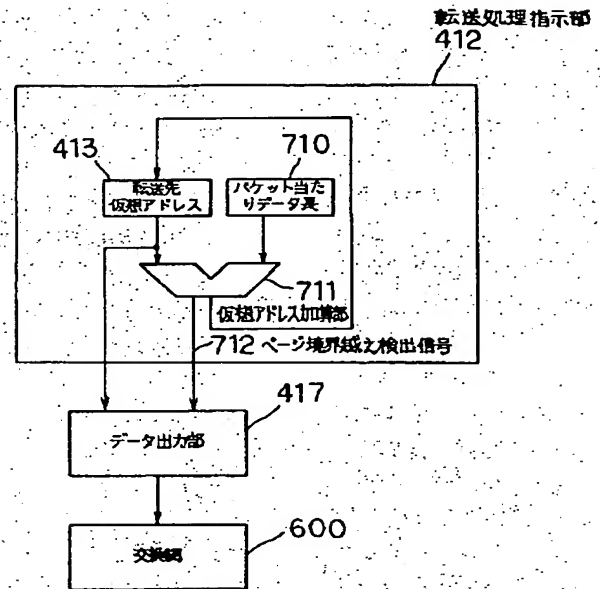
【図5】



【図6】



【図12】



【図8】

(a) リモート・ライト・パケット

命令項目	命令(Write)   データ長=4
ホスト項目	送信/受信計算機番号
アドレス項目	伝送先仮想アドレス
データ項目	伝送データ
データ項目	
データ項目	
データ項目	

(b) リモート・リード・パケット

命令項目	命令(Read)   データ長=1
ホスト項目	送信/受信計算機番号
アドレス項目	伝送先仮想アドレス
データ項目	伝送先仮想アドレス

(c) TLB先読みパケット

命令項目	命令(TLB)   データ長=0
ホスト項目	送信/受信計算機番号
アドレス項目	仮想アドレス

(d) 命令種別一覧

命令種別	説明
Write	リモート・ライト: 受信装置は伝送先仮想アドレスへデータ書き込み
Read	リモート・リード: 受信装置は伝送元仮想アドレスのデータを読み出し、ライト・パケットに伝送先仮想アドレスを設定し、送り元の送信装置に転送
TLB	仮想アドレスでTLBを参照し、TLBミスの場合はアドレス変換機から実アドレス・ベースを読み出し登録

(e) リモート・リード・パケット

命令項目	命令(Read)   データ長=1   PF
ホスト項目	送信/受信ホスト番号
アドレス項目	伝送元仮想アドレス
データ項目	伝送先仮想アドレス

\*PF: TLB先読み指示フラグ

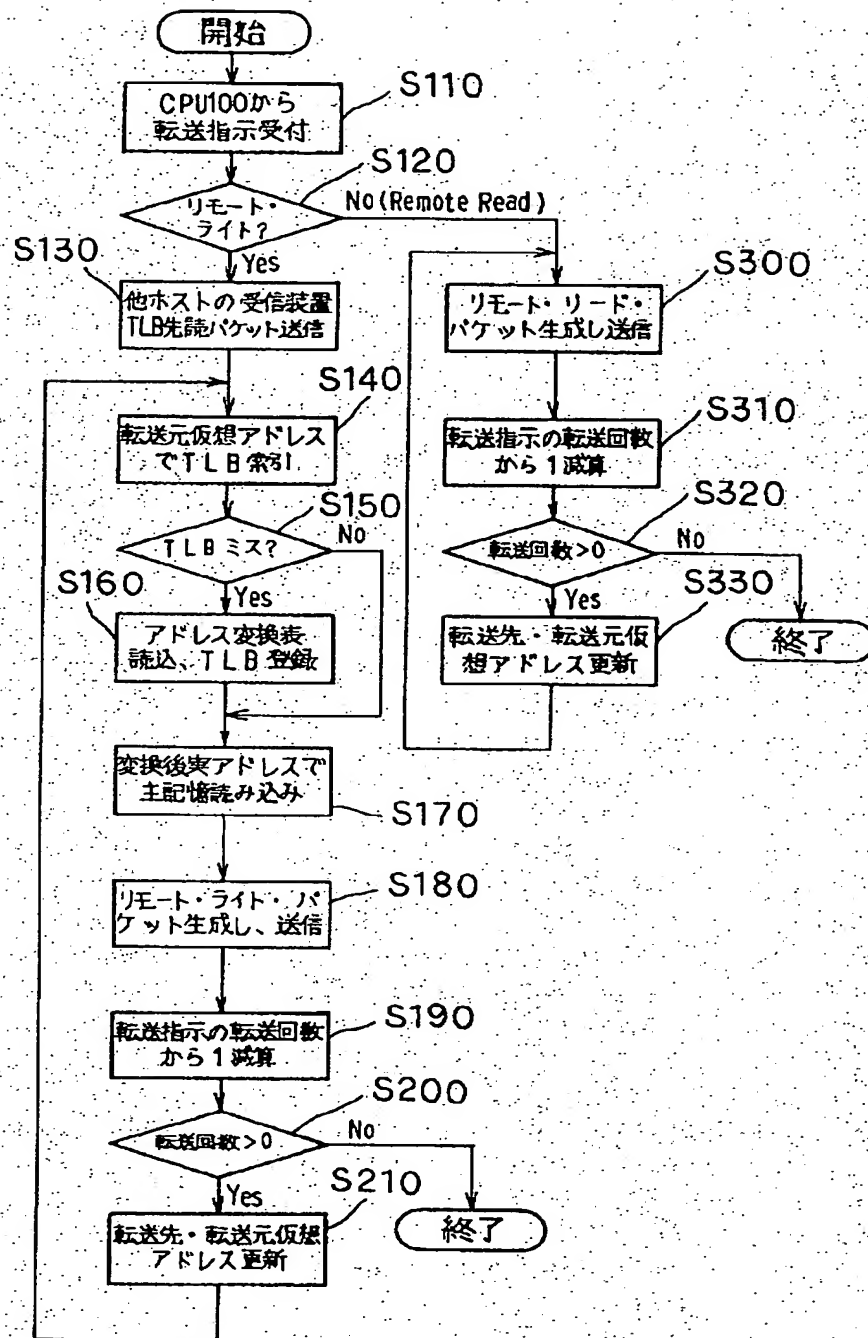
【図 7】

	計算機 1		計算機 2	
	CPU100	送信装置 410	受信装置 420	CPU100
順序 1	アドレス変換表 500 設定			アドレス変換表 500 設定
順序 2	リモート・ライト命令			
順序 3		転送先仮想アドレスを仮想アドレスとして含む TLB 先読みパケットを計算機 2 へ転送		
順序 4		転送元仮想アドレスで TLB を索引し転送元実アドレスに変換	TLB 先読みパケット内の仮想アドレスで TLB を索引し実アドレスに変換	
順序 5		TLB ミスの場合は、アドレス変換表 500 を用いて転送元実アドレスに変換	TLB ミスの場合は、アドレス変換表 500 を用いて実アドレスに変換	
順序 6		主記憶装置 300 の転送元実アドレスから書き込みデータを読み込む		
順序 7		書き込みデータと転送先仮想アドレスとを含むライト・パケットを計算機 2 へ転送		
順序 8			ライト・パケット内の転送先仮想アドレスで TLB を索引し転送先実アドレスに変換	
順序 9			TLB ミスの場合は、アドレス変換表 500 を用いて転送先実アドレスに変換	
順序 10			書き込みデータを主記憶装置 300 の転送先実アドレスに書き込む	

【図9】

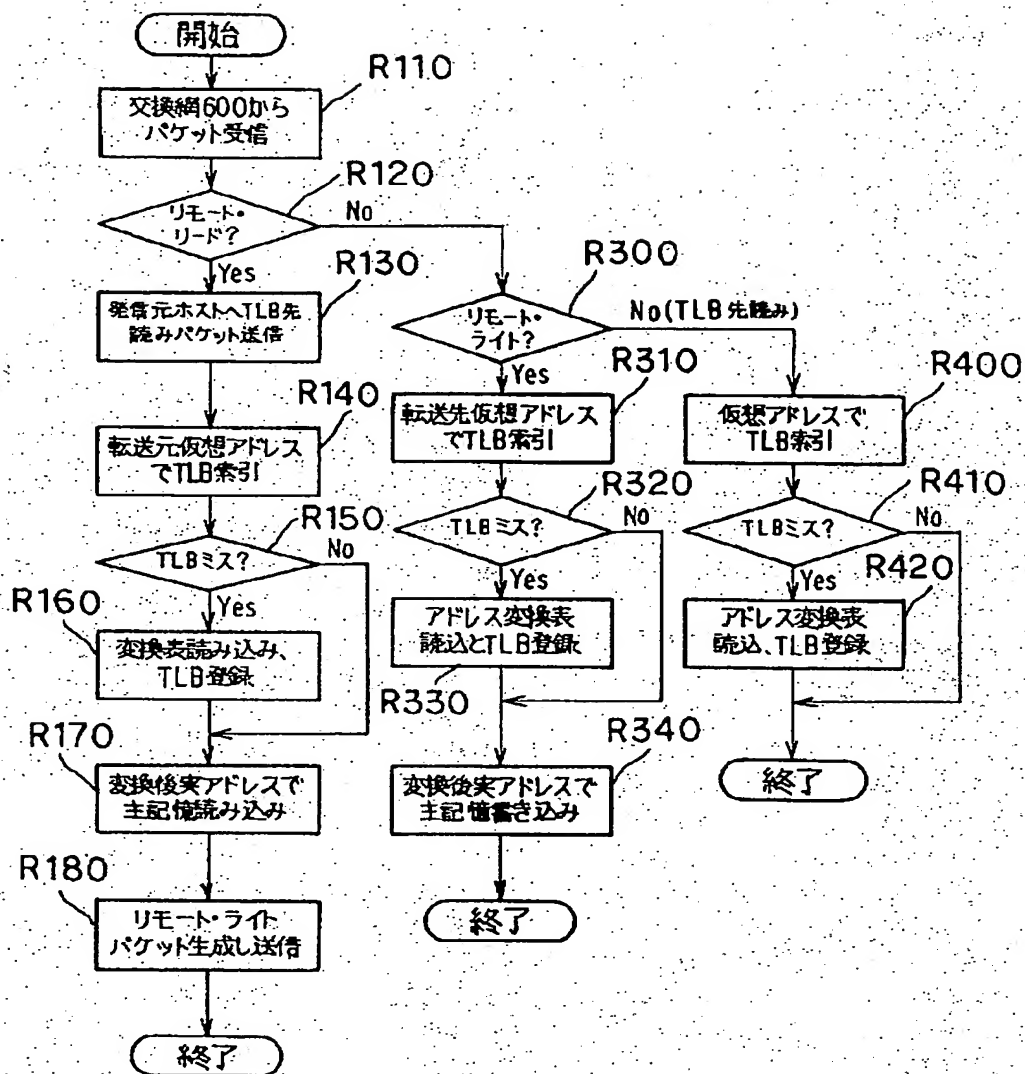
	計算機1			計算機2	
	CPU100	送信装置410	受信装置420	受信装置420	CPU100
順序1	アドレス変換表500設定				アドレス変換表500設定
順序2	リモート・リード命令				
順序3		転送元仮想アドレスと転送先仮想アドレスとを含むリード・パケットを計算機2へ転送			
順序4				リード・パケット内の転送先仮想アドレスを仮想アドレスとして含むTLB先読みパケットを計算機1へ転送	
順序5			TLB先読みパケット内の仮想アドレスでTLBを索引し、実アドレスに変換	リード・パケット内の転送元仮想アドレスでTLBを索引し、転送元実アドレスに変換	
順序6			TLBミスの場合は、アドレス変換表500を用いて実アドレスに変換	TLBミスの場合は、アドレス変換表600を用いて転送元実アドレスに変換	
順序7				主記憶装置300の転送元実アドレスから読み出しデータを読み出す	
順序8				読み出しデータとリード・パケット内の転送先論理アドレスとを含むライト・パケットを計算機1へ転送	
順序9			ライト・パケット内の転送先論理アドレスでTLBを索引し、転送先実アドレスに変換		
順序10			TLBミスの場合は、アドレス変換表500を用いて転送先実アドレスに変換		
順序11			読み出しデータを主記憶装置300の転送先実アドレスに書き込む		

【図10】

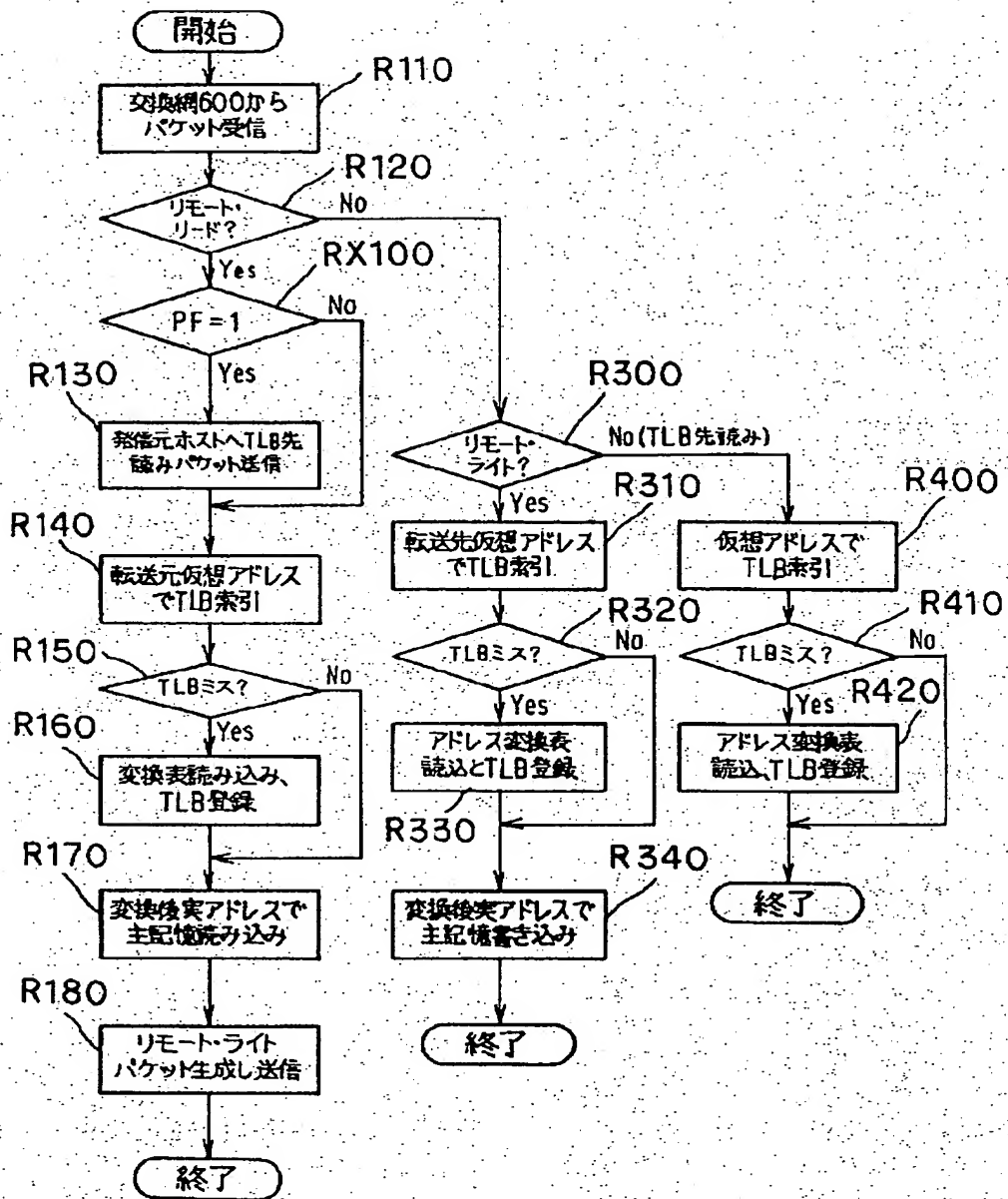




【図11】



【図13】



**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☒ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER: \_\_\_\_\_**

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**